



UNAP



FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA

**ESCUELA PROFESIONAL DE INGENIERÍA DE
SISTEMAS E INFORMÁTICA**

TESIS

**MÉTODOS ESTADÍSTICOS Y MODELOS DE MACHINE
LEARNING PARA LA PREDICCIÓN DE LA
RADIACIÓN SOLAR EN IQUITOS, 2023**

**PARA OPTAR EL TÍTULO PROFESIONAL DE
INGENIERO DE SISTEMAS E INFORMÁTICA**

PRESENTADO POR:

CÉSAR TIRADO TELLO

CHRISTIAN ALBERTO DULANTO TORRES

ASESOR:

Ing. CARLOS ALBERTO GARCIA CORTEGANO, Dr.

IQUITOS, PERÚ

2023

ACTA DE SUSTENTACIÓN



FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS E INFORMÁTICA

ACTA DE SUSTENTACIÓN DE TESIS - 2023

En Iquitos, en la modalidad presencial, a los 22 días del mes de diciembre del 2023, a horas 6:00 pm, se dio inicio a la sustentación de la Tesis Titulada: "MÉTODOS ESTADÍSTICOS Y MODELOS DE MACHINE LEARNING PARA LA PREDICCIÓN DE LA RADIACIÓN SOLAR EN IQUITOS 2023", presentado por los bachilleres: CHRISTIAN ALBERTO DULANTO TORRES y CÉSAR TIRADO TELLO, para optar el Título Profesional de Ingeniero de Sistemas e Informática, que otorga la Universidad de acuerdo a Ley y Estatuto.

El Jurado calificador y dictaminador designado mediante Resolución Decanal N°174-D-FISI-UNAP-2023, está integrado por:

- | | |
|--|------------|
| ✓ Ing. Alejandro Reátegui Pezo, Dr. | Presidente |
| ✓ Ing. Jimmy Max Ramírez Villacorta, Mgr. | Miembro |
| ✓ Ing. Ronald Percy Melchor Infantes, Mgr. | Miembro |



Luego de haber el Jurado escuchado con atención y formulado las preguntas necesarias, las cuales fueron respondidas: CORRECTAMENTE

El jurado después de las deliberaciones correspondientes, llegó a las siguientes conclusiones:

La Sustentación de la Tesis ha sido: APROBADO con la calificación de: 17

Estando los Bachilleres aptos para obtener el Título Profesional de **Ingeniero de Sistemas e Informática**.

Siendo las 19:00 horas se dio por terminado el acto de sustentación.

Ing. Alejandro Reátegui Pezo, Dr.
Presidente

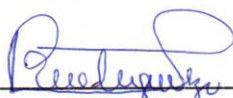
Ing. Jimmy Max Ramírez Villacorta, Mgr.
Miembro

Ing. Ronald Percy Melchor Infantes, Mgr.
Miembro

Ing. Carlos Alberto García Cortegano, Dr.
ASESOR

JURADO Y ASESOR

TESIS APROBADA EN SUSTENTACIÓN PÚBLICA EL DÍA **22** DE **DICIEMBRE**
AÑO **2023** EN LA FACULTAD DE INGENIERÍA DE SISTEMAS E
INFORMÁTICA DE LA UNIVERSIDAD NACIONAL DE LA AMAZONÍA
PERUANA, EN IQUITOS - PERÚ.



Ing. ALEJANDRO REÁTEGUI PEZO, Dr.

Presidente



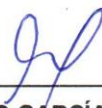
Ing. JIMMY MAX RAMÍREZ VILLACORTA, Mgr.

Miembro



Ing. RONALD PERCY MELCHOR INFANTES, Mgr.

Miembro



Ing. CARLOS ALBERTO GARCÍA CORTEGANO, Dr.

Asesor

RESULTADO DEL INFORME DE SIMILITUD

Reporte de similitud

NOMBRE DEL TRABAJO

FISI_TESIS_TIRADO TELLO_DULANTO TORRES.pdf

AUTOR

TIRADO TELLO / DULANTO TORRES

RECuento de palabras

8595 Words

RECuento de caracteres

48872 Characters

RECuento de páginas

35 Pages

Tamaño del archivo

909.4KB

Fecha de entrega

Feb 25, 2024 11:55 PM GMT-5

Fecha del informe

Feb 25, 2024 11:56 PM GMT-5

● 15% de similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para cada base de datos.

- 9% Base de datos de Internet
- Base de datos de Crossref
- 12% Base de datos de trabajos entregados
- 2% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossref

● Excluir del Reporte de Similitud

- Material bibliográfico
- Coincidencia baja (menos de 10 palabras)

Resumen

DEDICATORIA

Dedico este proyecto a mi amado hijo quien me motiva y me da fuerzas para seguir creciendo personal y profesionalmente.

Christian A. Dulanto Torres

Dedico este proyecto a mi esposa la cual me ha ayudado y motivado a seguir adelante en los momentos difíciles, a mis padres que han sabido formarme con buenos sentimientos y valores. También dedico a mis hijos quienes han sido mi mayor motivación para nunca rendirme y poder llegar a ser un ejemplo para ellos.

César Tirado Tello

AGRADECIMIENTO

A Dios nuestro creador por darnos la vida y darnos la oportunidad de formarnos profesionalmente con sólidos valores en esta prestigiosa universidad a la cual pertenecemos.

A nuestras familias por creer en nuestro potencial, por motivarnos a nunca rendirnos y siempre habernos apoyado incondicionalmente.

A nuestro asesor por compartir con nosotros sus conocimientos, tiempo y dedicación en la elaboración de esta tesis

ÍNDICE

| | |
|--|------------|
| PORTADA | i |
| ACTA DE SUSTENTACIÓN | ii |
| JURADO Y ASESOR | iii |
| RESULTADO DEL INFORME DE SIMILITUD | iv |
| DEDICATORIA | v |
| AGRADECIMIENTO | vi |
| ÍNDICE | vii |
| ÍNDICE DE TABLAS | ix |
| ÍNDICE DE FIGURAS | x |
| RESUMEN | xi |
| ABSTRACT | xii |
| INTRODUCCIÓN | 1 |
| CAPÍTULO I: MARCO TEÓRICO | 3 |
| 1.1. Antecedentes | 3 |
| 1.2. Bases teóricas | 6 |
| 1.2.1. Métodos de predicción. | 6 |
| 1.2.1.2.1. Modelo ARIMA. | 7 |
| 1.2.2. Radiación solar | 7 |
| 1.3. Definición de términos básicos | 9 |
| CAPÍTULO II: HIPÓTESIS Y VARIABLES | 12 |
| 2.1. Formulación de hipótesis. | 12 |
| 2.1.1. Hipótesis general. | 12 |
| 2.1.2. Hipótesis nula. | 12 |
| 2.2. Variables y operacionalización. | 12 |
| 2.2.1. Variables | 12 |
| 2.2.2. Operacionalización de las variables | 12 |
| CAPÍTULO III: METODOLOGÍA | 14 |
| 3.1. Diseño metodológico | 14 |
| 3.1.1. Tipo de investigación. | 14 |
| 3.1.2. Diseño de la investigación. | 14 |
| 3.2. Diseño muestral. | 14 |
| 3.3. Procedimientos de recolección de datos. | 16 |
| 3.4. Procesamiento y análisis de la información. | 17 |

| | |
|---|-----------|
| 3.5. Aspectos éticos. | 18 |
| CAPÍTULO IV: RESULTADOS | 20 |
| CAPÍTULO V: DISCUSIÓN | 28 |
| CAPÍTULO VI: CONCLUSIONES | 30 |
| CAPÍTULO VII: RECOMENDACIONES | 32 |
| CAPÍTULO VIII: REFERENCIAS BIBLIORÁFICAS | 34 |
| ANEXOS | 37 |
| Anexo 1: MATRIZ DE CONSISTENCIA | 38 |
| Anexo 2: DATA HISTORICA DE TEMPERATURA MEDIA(°C) DE LA CIUDAD DE IQUITOS. | 39 |
| Anexo 3: DATA HISTORICA DE HUMEDAD RELATIVA DE LA CIUDAD DE IQUITOS. | 40 |
| Anexo 4: DATA HISTORICA DE PRESION ATMOSFERICA DE LA CIUDAD DE IQUITOS. | 41 |
| Anexo 5: DATA HISTORICA DE VELOCIDAD DEL VIENTO DE LA CIUDAD DE IQUITOS. | 42 |
| Anexo 6: DATA HISTORICA DE RADIACION SOLAR DE LA CIUDAD DE IQUITOS. | 43 |
| Anexo 7: AJUSTE DEL MODELO ARIMA | 44 |
| Anexo 8: RADIACIÓN SOLAR DE IQUITOS | 45 |
| Anexo 9: GRÁFICOS DE MATLAB | 46 |
| Anexo 10: CÓDIGO FUENTE | 48 |

ÍNDICE DE TABLAS

| | |
|--|----|
| Tabla 1. Operacionalización de las variables..... | 13 |
| Tabla 2. Análisis de Desempeño del Modelo de Machine Learning (Redes Neuronales)..... | 20 |
| Tabla 3. Rendimiento del modelo estadístico ARIMA | 22 |
| Tabla 4. Comparación de Modelos | 24 |

ÍNDICE DE FIGURAS

| | |
|--|----|
| Figura 1. Radiación solar en Iquitos en función del tiempo. | 22 |
|--|----|

RESUMEN

Este proyecto investigativo desplegó una comparativa meticulosa entre dos métodos para predecir la radiación solar en Iquitos, utilizando el modelo estadístico ARIMA y un modelo de machine learning (ML) fundamentado en redes neuronales, con un enfoque mixto de diseño no experimental y datos suministrados por SENAMHI, englobando variables tales como radiación solar, humedad relativa y presión atmosférica. A pesar de que el modelo de redes neuronales exhibió un formidable rendimiento, evidenciando un 98.85% de explicación de la variabilidad en el conjunto de entrenamiento, se percibió una disminución de su eficiencia en las fases de validación y prueba, sugiriendo un sobreajuste potencial. Contrariamente, el modelo ARIMA, aunque presentó un rendimiento simplemente aceptable con un MSE de 45.651 y un R^2 de 0.8500, demostró una robustez constante. La comparativa explicitó que, pese a que las redes neuronales ostentan una superioridad en cuanto a capacidad predictiva, explicando un 95.51% de la variabilidad y exhibiendo un menor error cuadrático medio, es imperativo sopesar otros elementos críticos como la interpretación del modelo y los recursos computacionales al culminar decisiones. En definitiva, mientras que los métodos de ML, particularmente las redes neuronales, demostraron ser particularmente perspicaces en la predicción de la radiación solar, la exploración y optimización de diferentes enfoques modelísticos se presenta como un horizonte esencial para futuras investigaciones en la materia.

Palabras clave: Radiación solar, Iquitos, modelo ARIMA, redes neuronales, machine learning (ML), predicción, comparación, precisión, confiabilidad.

ABSTRACT

This investigative project deployed a meticulous comparison between two methods for predicting solar radiation in Iquitos, utilizing the statistical ARIMA model and a machine learning model grounded in neural networks, with a mixed non-experimental design approach and data provided by SENAMHI, encompassing variables such as solar radiation, relative humidity, and atmospheric pressure. Despite the neural network model displaying formidable performance, accounting for 98.85% of the variability in the training dataset, a decline in its efficiency was observed in the validation and testing phases, hinting at potential overfitting. Conversely, the ARIMA model, while showcasing merely acceptable performance with an MSE of 45.651 and an R^2 of 0.8500, demonstrated steady robustness. The comparison elucidated that, although neural networks possess superiority in predictive capacity, explaining 95.51% of the variability and showcasing a lower mean squared error, it is imperative to weigh other critical elements such as model interpretation and computational resources when finalizing decisions. Ultimately, while machine learning methods, particularly neural networks, proved to be especially insightful in predicting solar radiation, the exploration and optimization of different modeling approaches emerge as an essential horizon for future research in the field.

Keywords: Solar radiation, Iquitos, ARIMA model, neural networks, machine learning, prediction, comparison, accuracy, reliability.

INTRODUCCIÓN

En el contemporáneo escenario marcado por el cambio climático y una ascendente inquietud medioambiental, la radiación solar se erige como un pilar energético renovable y sostenible de magnitud innegable. Su potencial para mitigar la dependencia de los combustibles fósiles y decelerar las emisiones de gases de efecto invernadero ha sido profundamente estudiado, ilustrando un panorama auspicioso en el campo energético (IPCC, 2021). Es en esta coyuntura que la exactitud en la predicción de radiación solar se convierte en un elemento vital para la operativa eficaz y progresión de sistemas energéticos solares, especialmente en zonas con climatologías tan complejas y variantes como Iquitos.

Ubicada en la porción amazónica de Perú, Iquitos surge como un enclave de crucial relevancia en la investigación solar, debido a su climatología diversa y fenómenos atmosféricos particulares. A pesar de la trascendencia de la región, prevalece una notoria carestía de estudios que enfrenten métodos estadísticos y de machine learning en la predicción solar en el área, desencadenando un desafío en la selección de técnicas optimizadas para tal fin. Así, este estudio pretende desentrañar y evaluar estos enfoques en Iquitos, con el objetivo de discernir cuál se erige con mayor precisión y confiabilidad en las predicciones.

La problemática que orienta esta investigación se manifiesta de la siguiente manera: ¿Qué método, ya sea estadístico o de machine learning, proporciona proyecciones de radiación solar en Iquitos para 2023 que se caractericen por su precisión y confiabilidad? Con el propósito cardinal de cotejar un método estadístico con un modelo de machine learning en la predicción de radiación solar para Iquitos en el aludido año, se articulan objetivos específicos que abarcan la evaluación de la exactitud y fiabilidad de ambos modelos y una comparación meticulosa entre ambos.

Este estudio se distingue por su relevancia en cuanto al avance del saber científico en la esfera de la energía solar y en sus implicancias prácticas para la conceptualización y planificación de iniciativas de energía solar en la región amazónica. Además, la indagación ostenta un valor teórico y metodológico al

contraponer enfoques predictivos y evaluar indicadores de desempeño, erigiéndose como un referente para futuras exploraciones en diversos contextos geográficos y climáticos.

Desde la perspectiva metodológica, se adopta un enfoque mixto y aplicado, orientado a solventar problemáticas concretas y escudriñar las relaciones entre variables. El diseño no experimental y transversal facilita el análisis de datos de radiación solar y variables meteorológicas recabadas durante un período establecido en Iquitos. La población investigada engloba la totalidad de registros meteorológicos en la región, empleando un muestreo no probabilístico y capitalizando el 100% de los datos accesibles.

Con la realización de este estudio, se aspira a imprimir un aporte significativo al campo de la predicción de radiación solar, suministrando modelos de gran valor para la instauración de sistemas de energía solar en Iquitos y fomentando la adopción de energías limpias y sostenibles en la región, en favor del ambiente y la colectividad. La estructura de la tesis se desplegará siguiendo una progresión lógica, introduciendo la revisión literaria, la metodología, los resultados y, ulteriormente, las conclusiones y recomendaciones.

La exploración se estructura de manera detallada y organizada para facilitar la comprensión del lector acerca de los distintos componentes y descubrimientos. El primer capítulo proporciona un marco teórico robusto que sienta las bases para el estudio, seguido por el segundo capítulo, donde se presentan la hipótesis central y las variables en juego. El tercer capítulo está dedicado a exponer la metodología adoptada para llevar a cabo la investigación, ofreciendo detalles sobre los procedimientos y técnicas utilizadas. Posteriormente, en el cuarto capítulo, se revelan los resultados obtenidos, los cuales son analizados y debatidos en profundidad en el capítulo quinto bajo el título de discusiones. La sexta sección del trabajo recoge las conclusiones derivadas de la investigación, destacando los hallazgos clave y su significado. En el séptimo capítulo, se ofrecen recomendaciones basadas en los resultados y conclusiones del estudio, y finalmente, el octavo capítulo recopila la bibliografía y referencias consultadas durante la investigación.

CAPÍTULO I: MARCO TEÓRICO

1.1. Antecedentes

En este estudio, Li et al. (2018) tuvieron como objetivo desarrollar un modelo de predicción de radiación solar basado en redes neuronales recurrentes LSTM (Long Short-Term Memory). La investigación utilizó un enfoque cuantitativo y un diseño de investigación experimental. Los datos fueron recopilados de una estación meteorológica en China durante el período de 2013 a 2016. La muestra estuvo compuesta por series temporales de radiación solar, temperatura, humedad y velocidad del viento. Para la recolección y procesamiento de datos, se utilizó la biblioteca Keras en Python. El análisis de datos incluyó la evaluación del rendimiento del modelo LSTM mediante métricas como el error cuadrático medio (RMSE) y el coeficiente de determinación (R^2). Los resultados mostraron que el modelo LSTM logró un rendimiento superior en comparación con otros métodos de predicción de radiación solar. En conclusión, los autores sugirieron que las redes neuronales LSTM podrían ser una herramienta valiosa para la predicción precisa de la radiación solar.

Gensollen y Ortega (2019) investigaron la aplicación de técnicas de machine learning en la predicción de radiación solar. El objetivo de este estudio cuantitativo y experimental fue comparar diferentes modelos de machine learning, como Support Vector Machines (SVM), Decision Trees, Random Forest y Gradient Boosting. La población del estudio consistió en datos de radiación solar global obtenidos de la base de datos del National Solar Radiation Database (NSRDB) en los Estados Unidos para el período 1998-2017. La recolección y procesamiento de datos se realizó utilizando el lenguaje de programación Python y bibliotecas como Scikit-learn y Pandas. El análisis de datos se centró en evaluar el rendimiento de los diferentes modelos de machine learning mediante métricas como el RMSE y el coeficiente de correlación de Pearson. Los resultados indicaron que el modelo Random Forest mostró el mejor rendimiento en la predicción de radiación solar. En conclusión, los autores destacaron la

importancia de explorar técnicas de machine learning para mejorar la precisión en la predicción de radiación solar.

En este estudio, Malik y Bilal (2020) desarrollaron un modelo de predicción de radiación solar horaria utilizando técnicas de deep learning. El objetivo de la investigación fue evaluar la efectividad de las redes neuronales convolucionales (CNN) y las redes neuronales recurrentes (RNN) en la predicción de radiación solar. La investigación adoptó un enfoque cuantitativo y un diseño experimental. La población del estudio consistió en datos de radiación solar horaria de Islamabad, Pakistán, durante el período de 2005 a 2015. La muestra incluyó mediciones de radiación solar, temperatura, humedad y velocidad del viento. La recolección y procesamiento de datos se realizó utilizando el lenguaje de programación Python y bibliotecas como TensorFlow y Keras. El análisis de datos se centró en comparar el rendimiento de los modelos CNN y RNN mediante métricas como el RMSE y el coeficiente de determinación (R^2). Los resultados mostraron que las RNN, en particular las redes LSTM, ofrecían un mejor rendimiento en la predicción de radiación solar horaria en comparación con las CNN. En conclusión, los autores destacaron el potencial de las técnicas de deep learning, especialmente las RNN, para mejorar la precisión en la predicción de radiación solar horaria.

En esta revisión sistemática, Zhao y Wei (2020) evaluaron la literatura sobre la predicción de radiación solar utilizando técnicas de machine learning y deep learning. El objetivo del estudio fue identificar los avances y desafíos en el uso de estos enfoques para la predicción de radiación solar. Los autores utilizaron una metodología de revisión sistemática para seleccionar y analizar estudios relevantes publicados entre 2010 y 2019. La población del estudio incluyó investigaciones que utilizaron modelos de machine learning y deep learning en la predicción de radiación solar. La muestra consistió en 68 estudios seleccionados según criterios de inclusión y exclusión. El análisis de datos incluyó la identificación de tendencias, enfoques y desafíos en la literatura. Los resultados mostraron que los modelos de machine learning y deep learning,

como las redes neuronales artificiales, SVM y LSTM, han demostrado un buen rendimiento en la predicción de radiación solar. Sin embargo, los autores también identificaron desafíos en términos de la calidad y cantidad de datos, así como la selección de características y el ajuste de hiperparámetros. En conclusión, Zhao y Wei (2020) destacaron la necesidad de continuar explorando y mejorando los enfoques de machine learning y deep learning para la predicción de radiación solar.

En este estudio, Gutiérrez-Madroñal et al. (2021) propusieron un enfoque para la predicción de radiación solar utilizando técnicas de machine learning y optimización metaheurística. El objetivo de la investigación fue mejorar la precisión de los modelos de machine learning mediante la optimización de hiperparámetros utilizando algoritmos metaheurísticos. La investigación adoptó un enfoque cuantitativo y un diseño experimental. La población del estudio consistió en datos de radiación solar global obtenidos de estaciones meteorológicas en España durante el período 2004-2018. La muestra incluyó mediciones de radiación solar, temperatura, humedad y velocidad del viento. La recolección y procesamiento de datos se realizó utilizando el lenguaje de programación Python y bibliotecas como Scikit-learn y Pandas. El análisis de datos se centró en evaluar el rendimiento de los modelos de machine learning optimizados mediante algoritmos metaheurísticos, como la optimización por enjambre de partículas (PSO) y el algoritmo genético (GA). Se emplearon métricas como el RMSE y el coeficiente de determinación (R^2) para evaluar el rendimiento de los modelos. Los resultados mostraron que la optimización de hiperparámetros utilizando técnicas metaheurísticas mejoró significativamente la precisión de los modelos de machine learning en la predicción de radiación solar. En conclusión, los autores destacaron el potencial de combinar machine learning y optimización metaheurística para mejorar la precisión en la predicción de radiación solar y recomendaron investigaciones futuras para explorar otros algoritmos de optimización y técnicas de machine learning.

1.2. Bases teóricas

1.2.1. Métodos de predicción.

1.2.1.1. Modelo de Machine Learning(ML)

El ML es una rama de la inteligencia artificial que se centra en el desarrollo de algoritmos y modelos que pueden aprender de los datos para mejorar su rendimiento en tareas específicas (Mitchell, 1997). Los modelos de machine Learning pueden ser supervisados, no supervisados o por refuerzo, dependiendo de cómo se utilizan los datos de entrada y salida durante el proceso de aprendizaje (Bishop, 2006). Algunos de los algoritmos de machine learning más comunes utilizados en la predicción de radiación solar incluyen redes neuronales artificiales (ANN), máquinas de vectores de soporte (SVM), árboles de decisión, bosques aleatorios y métodos de impulso (Gensollen & Ortega, 2019).

El deep learning es una subárea del machine learning que se centra en el desarrollo de redes neuronales artificiales con múltiples capas ocultas, lo que permite aprender representaciones jerárquicas de datos (LeCun et al., 2015). Las redes neuronales convolucionales (CNN) y las redes neuronales recurrentes (RNN) son dos tipos de arquitecturas de deep learning comúnmente utilizadas en la predicción de radiación solar (Malik & Bilal, 2020). Las CNN son especialmente efectivas para aprender características locales en datos espaciales, como imágenes, mientras que las RNN, en particular las redes LSTM (Long Short-Term Memory), son adecuadas para aprender dependencias temporales en series temporales de datos (Goodfellow et al., 2016).

La optimización de hiperparámetros es un proceso importante en el desarrollo de modelos de machine learning y deep learning, ya que afecta directamente el rendimiento y la precisión del modelo. Los hiperparámetros son valores configurables que determinan la estructura y el comportamiento de un algoritmo durante el aprendizaje y la inferencia (Bergstra & Bengio, 2012). La optimización de hiperparámetros puede realizarse mediante enfoques de búsqueda en cuadrícula, búsqueda aleatoria o algoritmos metaheurísticos, como la

optimización por enjambre de partículas (PSO) y el algoritmo genético (GA) (Gutiérrez-Madroñal et al., 2021).

1.2.1.2. Modelos estadísticos.

Los modelos estadísticos son representaciones matemáticas que describen la relación entre variables y permiten realizar predicciones y estimaciones. Estos modelos se basan en la teoría de la probabilidad y las estadísticas, y son ampliamente utilizados en diversas disciplinas como la economía, la meteorología y la biología (Pindyck & Rubinfeld, 1998).

1.2.1.2.1. Modelo ARIMA.

El modelo autorregresivo integrado de media móvil (ARIMA, por sus siglas en inglés) es un enfoque estadístico utilizado para predecir series temporales, analizando y modelando las autocorrelaciones observadas en los datos. Los modelos ARIMA combinan tres componentes: autorregresión (AR), integración (I) y media móvil (MA).

La autorregresión se refiere a la relación entre una observación actual y sus observaciones anteriores, la integración implica la diferenciación de la serie temporal para hacerla estacionaria, y la media móvil describe la relación entre una observación actual y una combinación lineal de errores anteriores (Box et al., 2013).

Los modelos ARIMA son útiles para predecir series temporales con tendencias y patrones estacionales. Estos modelos se han utilizado ampliamente en la predicción de variables económicas, financieras, climáticas y de otra índole. Los modelos ARIMA pueden ajustarse a distintos grados de complejidad y flexibilidad, dependiendo de la naturaleza de los datos y de los objetivos de predicción (Hyndman & Athanasopoulos, 2018).

1.2.2. Radiación solar

La radiación solar es la energía electromagnética emitida por el Sol que llega a la Tierra en forma de ondas electromagnéticas, siendo esencial

para mantener la vida en el planeta y para la generación de energía a través de sistemas fotovoltaicos y térmicos solares (Iqbal, 2012). La radiación solar se compone principalmente de radiación directa y radiación difusa. La radiación directa es la que llega directamente del Sol sin cambios en su trayectoria, mientras que la radiación difusa es la que ha sido dispersada en múltiples direcciones por la atmósfera antes de llegar a la superficie terrestre (Duffie & Beckman, 2013).

1.2.2.1. Predicción de radiación solar

La predicción de la radiación solar es crucial para la planificación y operación eficiente de sistemas de energía solar y para la gestión de la demanda energética. Existen varios enfoques y métodos para predecir la radiación solar, que pueden clasificarse en modelos empíricos, modelos físicos y modelos basados en datos (Paoli et al., 2010). Los modelos empíricos se basan en relaciones matemáticas y estadísticas derivadas de datos históricos, mientras que los modelos físicos describen los procesos atmosféricos que afectan la radiación solar. Los modelos basados en datos, como los modelos de machine learning, utilizan algoritmos para aprender patrones en datos históricos y predecir valores futuros (Sharma et al., 2020).

1.2.2.1.1. Precisión y Confiabilidad en las Predicciones de Radiación Solar

Precisión

La precisión se refiere a cuán cerca están las predicciones de un modelo de los valores verdaderos. En el contexto de la predicción de radiación solar, la precisión se preocupa por cuán exactamente se predice la radiación solar en comparación con las medidas reales (Wilks, 2011).

Confiabilidad

La confiabilidad, en cambio, se refiere a la consistencia de un modelo en sus predicciones a lo largo del tiempo y en diferentes condiciones. Un modelo confiable ofrecerá resultados similares en

diferentes momentos y condiciones, independientemente de su precisión (Jolliffe and Stephenson, 2012).

1.2.2.1.2. Diferencia entre Precisión y Confiabilidad

Enfoque: La precisión se enfoca en la cercanía de la predicción a los valores verdaderos, mientras que la confiabilidad se enfoca en la repetibilidad y consistencia de las predicciones (Wilks, 2011).

Relación: Un modelo puede ser preciso sin ser confiable y viceversa. Por ejemplo, un modelo que funcione excepcionalmente bien en un conjunto de datos específico puede no ser confiable si no funciona igual de bien en otros conjuntos de datos (Jolliffe and Stephenson, 2012).

Importancia en la Predicción de Radiación Solar: La precisión es vital para asegurar que las predicciones de radiación solar estén en línea con los valores reales, mientras que la confiabilidad asegura que las predicciones sean consistentes a lo largo del tiempo, lo cual es esencial para la planificación a largo plazo y la eficiencia energética (Zhang et al., 2016).

1.3. Definición de términos básicos

ARIMA: Modelo autorregresivo integrado de media móvil (ARIMA, por sus siglas en inglés) es un enfoque estadístico utilizado para predecir series temporales, analizando y modelando las autocorrelaciones observadas en los datos (Box, Jenkins & Reinsel, 2013).

Coeficiente de determinación (R^2): El coeficiente de determinación es una medida que indica qué proporción de la variabilidad observada en los datos puede ser explicada por el modelo de predicción, con valores que oscilan entre 0 y 1 (Willmott & Matsuura, 2005).

Deep learning: El deep learning es una subárea del machine learning que se centra en el desarrollo de redes neuronales artificiales con múltiples capas ocultas, lo que permite aprender representaciones jerárquicas de datos (LeCun et al., 2015).

Error absoluto medio (MAE): El error absoluto medio es una métrica que cuantifica la diferencia promedio entre los valores predichos y los valores observados, sin tener en cuenta la dirección del error (Willmott & Matsuura, 2005).

Error cuadrático medio (MSE): El error cuadrático medio es una métrica que cuantifica la diferencia promedio entre los valores predichos y los valores observados, elevando al cuadrado las diferencias individuales para penalizar los errores grandes (Willmott & Matsuura, 2005).

Humedad relativa: La humedad relativa es una medida de la cantidad de vapor de agua presente en el aire, expresada como un porcentaje del máximo posible de vapor de agua que el aire puede contener a una temperatura específica (AMS, 2021).

Machine learning: El machine learning es una rama de la inteligencia artificial que se centra en el desarrollo de algoritmos y modelos que pueden aprender de los datos para mejorar su rendimiento en tareas específicas (Mitchell, 1997).

MATLAB: MATLAB es un entorno de programación y lenguaje de alto nivel utilizado para realizar cálculos numéricos, análisis de datos, desarrollo de algoritmos y visualización (MathWorks, 2021).

Precipitaciones: Las precipitaciones son todas las formas de agua, ya sea líquida o sólida, que caen del cielo y llegan a la superficie terrestre, como lluvia, nieve, granizo, etc. Se mide en milímetros (mm) (AMS, 2021).

Predicción: Es la estimación o proyección de un evento futuro basada en información y modelos existentes (Hyndman and Athanasopoulos, 2018).

Predicción Precisa: Se refiere a la cercanía de la predicción a los valores verdaderos o reales, reflejando la exactitud de la estimación (Wilks, 2011).

Predicción Confiable: Indica la consistencia y repetibilidad de las predicciones a lo largo del tiempo y bajo diferentes condiciones, sin necesariamente referirse a su cercanía a los valores verdaderos (Jolliffe and Stephenson, 2012).

Presión atmosférica: La presión atmosférica es la fuerza por unidad de área ejercida por el peso del aire sobre una superficie y se mide en hectopascales (hPa) (AMS, 2021).

Radiación solar: La radiación solar es la energía electromagnética emitida por el Sol que llega a la Tierra en forma de ondas electromagnéticas. Esta energía es esencial para mantener la vida en el planeta y para la generación de energía a través de sistemas fotovoltaicos y térmicos solares (Iqbal, 2012).

Redes neuronales artificiales (ANN): Las ANN son modelos matemáticos inspirados en el funcionamiento de las neuronas biológicas y las redes neuronales del cerebro. Estos modelos son capaces de aprender y adaptarse a partir de datos de entrada y salida para realizar tareas de clasificación, predicción y aproximación de funciones (Bishop, 2006).

Temperatura ambiental: La temperatura ambiental es la medida del grado de calor o frío del aire en un lugar específico y se expresa en grados Celsius (°C) (AMS, 2021).

Velocidad del viento: La velocidad del viento es una medida de la rapidez con la que se mueve el aire en la atmósfera y se expresa en metros por segundo (m/s) (AMS, 2021).

CAPÍTULO II: HIPOTESIS Y VARIABLES

2.1. Formulación de hipótesis.

2.1.1. Hipótesis general.

Hi: Existe una diferencia significativa en la precisión y confiabilidad de las predicciones de la radiación solar entre el método estadístico y el modelo de machine Learning en Iquitos 2023.

2.1.2. Hipótesis nula.

H₀: No existe diferencia significativa en la precisión y confiabilidad de las predicciones de la radiación solar entre el método estadístico y el modelo de machine Learning en Iquitos para el año 2023.

2.2. Variables y operacionalización.

2.2.1. Variables

Variable independiente: Método de predicción (modelo basado en técnicas de machine learning o modelo estadístico tradicional).

Variable dependiente: Precisión y confiabilidad en las predicciones de radiación solar en Iquitos.

2.2.2. Operacionalización de las variables

Tabla 1. Operacionalización de las variables.

| Variable | Definición | Tipo | Indicadores | Escala de medición | Categoría | Medio de verificación |
|---|--|--------------|-------------------------------------|--------------------|-----------------------------|---|
| Variable independiente : Método de predicción | Se refiere al enfoque utilizado para predecir la radiación solar en Iquitos, ya sea un modelo basado en técnicas de machine Learning o un modelo estadístico tradicional | cualitativa | 1. Modelo de aprendizaje automático | Nominal | Redes neuronales | Evaluación de rendimiento usando métricas como RMSE, MAE, en conjuntos de datos de prueba. |
| | | | 2. Modelo estadístico tradicional | Nominal | ARIMA | Validación cruzada, análisis de residuos. |
| Variable dependiente: Precisión y confiabilidad en las predicciones de radiación solar | La exactitud y consistencia en las predicciones de radiación solar en Iquitos generadas por los modelos de predicción. | Cuantitativa | 1. Precisión de la predicción. | Intervalo | -Alta -Media -- Baja. | 1. Error cuadrático medio (MSE) 2. Coeficiente de correlación 3. Coeficiente de determinación (R^2) |
| | | | 2. Confiabilidad de la predicción | Intervalo | -Alta -Media -- Baja. | 1. Coeficiente de correlación 2. Índice de concordancia |

Fuente: Elaboración propia

CAPÍTULO III: METODOLOGÍA

3.1. Diseño metodológico

3.1.1. Tipo de investigación.

La investigación adoptó un enfoque mixto, que combinó métodos cualitativos y cuantitativos, proporcionando una visión más completa del estudio. En este escenario, los métodos cualitativos contribuyeron a comprender los procesos y métodos de predicción, mientras que los cuantitativos permitieron medir la precisión y confiabilidad (Creswell, J.W., 2014).

Se trató de una investigación aplicada, centrada en resolver un problema práctico más que en simplemente contribuir al cuerpo teórico del conocimiento. La aplicación concreta fue la mejora de las predicciones de la radiación solar, la cual cuenta con aplicaciones prácticas en campos como la energía solar y la meteorología (Sekaran, U., and Bougie, R., 2016).

El nivel del estudio es comparativo-predictivo, porque comparó métodos de machine learning y estadísticos para predecir la radiación solar en Iquitos en 2023.

3.1.2. Diseño de la investigación.

En cuanto al diseño de investigación, se optó por un enfoque no experimental y transversal, ya que no se manipularon las variables ni se efectuó un seguimiento a lo largo del tiempo. Se analizaron los datos de radiación solar y otras variables meteorológicas que se recopilaron en un periodo determinado.

3.2. Diseño muestral.

La población de estudio abarcó la totalidad de los registros vinculados con datos meteorológicos en Iquitos, incorporando variables como la radiación solar, humedad relativa, temperatura ambiente, presión atmosférica, y velocidad del viento. Estos datos fueron adquiridos manualmente de una base de datos en Excel suministrada por el Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI).

Dada la disponibilidad y calidad de los datos, la muestra incluyó el 100% de la población, representando así un conjunto de datos exhaustivo. Los registros meteorológicos históricos comprendieron un periodo de 96 meses (de enero de 2015 a diciembre de 2022), estructurados de la siguiente manera:

72 meses para el entrenamiento (75% del total),

14 meses destinados a la validación (15% del total),

10 meses reservados para la prueba (10% del total).

Tamaño de Muestra: En la inteligencia artificial y, en particular, en el aprendizaje automático, se prefiere utilizar todo el conjunto de datos disponible (100% de la población en este caso) si es manejable. Esto se justifica ya que más datos suelen proporcionar una mejor representación de la población subyacente y permiten entrenar modelos más precisos y robustos. Por lo tanto, no fue necesario calcular el tamaño de la muestra en el sentido tradicional.

Nivel de Significancia: El nivel de significancia es un concepto que se utiliza en la inferencia estadística para controlar la tasa de error Tipo I. En el contexto de la inteligencia artificial, este concepto pudo no ser aplicable directamente, ya que los modelos fueron evaluados usando métricas como RMSE, MAE, etc., en lugar de pruebas de hipótesis.

El muestreo utilizado en este proyecto se consideró no probabilístico, ya que se utilizó todo el conjunto de datos disponible, en lugar de seleccionar una muestra aleatoria. En este caso, no hubo elementos de aleatoriedad en la selección de la muestra.

Criterios de Inclusión y Exclusión

Criterios de Inclusión: Se incluyeron todos los datos de radiación solar y variables meteorológicas relacionadas registradas en Iquitos desde enero de 2015 hasta diciembre de 2022.

Criterios de Exclusión: Se excluyeron aquellos datos que estaban incompletos, eran inexactos o resultaban irrelevantes para la predicción

de la radiación solar. La calidad de los datos fue fundamental para entrenar modelos precisos y confiables.

3.3. Procedimientos de recolección de datos.

Técnicas:

Los datos fueron recolectados manualmente de la base de datos en Excel proporcionada por el Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI). Estos incluyeron registros históricos mensualizados de enero a diciembre de radiación solar, humedad relativa, temperatura ambiente, presión atmosférica y velocidad del viento en Iquitos durante el período de enero de 2015 a diciembre de 2022.

La selección y manejo de los datos se realizó de manera que reflejara los objetivos del estudio y las características de las unidades elementales de estudio, garantizando así que los datos fueran relevantes y útiles para la investigación propuesta.

Instrumentos:

Instrumento de Recopilación de Datos: Base de Datos en Excel

Tipo: Documento electrónico.

Descripción: Contenía registros históricos mensuales, desde enero de 2015 hasta diciembre de 2022, de diversas variables meteorológicas como la radiación solar, humedad relativa, temperatura ambiente, presión atmosférica y velocidad del viento en Iquitos.

Validación: La veracidad de esta base de datos se garantizó mediante una revisión detallada, asegurando que los datos fueran coherentes, completos y precisos. La autenticidad se confirmó por la fuente confiable, en este caso, el SENAMHI. Este paso fue esencial para garantizar que los datos estuvieran en línea con los objetivos del estudio.

Instrumento de Procesamiento y Análisis de Datos: Software MATLAB

Tipo: Herramientas tecnológicas.

Descripción: MATLAB fue el software seleccionado para preprocesar los datos meteorológicos, entrenar los modelos predictivos y evaluar su rendimiento. Su elección se basó en su compatibilidad con los métodos de modelado y análisis estadístico que se emplearon.

Validación: La validación de MATLAB se centró en la compatibilidad con los métodos seleccionados y en la actualización de las bibliotecas y funciones utilizadas. Esto garantizó que el software funcionara correctamente y estuviera alineado con las metodologías propuestas en el estudio.

3.4. Procesamiento y análisis de la información.

El procesamiento y análisis de datos para este proyecto de investigación sobre la predicción de radiación solar en Iquitos se llevó a cabo de manera secuencial, siguiendo los pasos detallados a continuación:

1. Revisión y verificación de la calidad de los datos: Se examinaron los datos recopilados y preprocesados para asegurarse de que estuvieran completos, consistentes y libres de errores antes de comenzar el análisis.
2. Selección de variables y características: Se identificaron las variables y características relevantes para la construcción de los modelos de predicción, considerando su relación con la radiación solar y su potencial contribución a la precisión de las predicciones.
3. Desarrollo del modelo estadístico (ARIMA): Se seleccionó el enfoque estadístico ARIMA (modelo autorregresivo integrado de media móvil) como método estadístico adecuado y se construyó el modelo utilizando el conjunto de datos de entrenamiento. Se ajustaron y optimizaron los parámetros del modelo para mejorar su desempeño en la predicción de la radiación solar.
4. Desarrollo del modelo de machine learning (Redes Neuronales): Se seleccionó un algoritmo de machine learning basado en Redes Neuronales apropiado y se construyó el modelo utilizando el conjunto

de datos de entrenamiento. Al igual que con el modelo ARIMA, se ajustaron y optimizaron los hiperparámetros del modelo de Redes Neuronales para mejorar su desempeño en la predicción de la radiación solar.

5. Evaluación de los modelos: Ambos modelos, ARIMA y Redes Neuronales, fueron evaluados utilizando el conjunto de datos de validación y métricas relevantes, coeficiente de correlación (R), el error cuadrático medio (MSE) y el coeficiente de determinación (R^2).
6. Comparación de los modelos: Los resultados obtenidos con ambos modelos, ARIMA y Redes Neuronales, se compararon en términos de precisión y confiabilidad de las predicciones. Este análisis permitió determinar cuál de los dos enfoques, estadístico o machine learning, ofrecía mejores resultados en la predicción de la radiación solar en Iquitos.
7. Interpretación de los resultados: Se analizaron los resultados de la comparación, considerando posibles explicaciones y factores que pudieron haber influido en los desempeños de los modelos ARIMA y Redes Neuronales. También se discutieron las implicaciones de los resultados en términos de aplicaciones prácticas y futuras investigaciones.

Siguiendo esta secuencia de pasos, se procesaron y analizaron los datos para llevar a cabo la comparación de los modelos de predicción de radiación solar en Iquitos, utilizando el modelo estadístico ARIMA y el modelo de machine learning basado en Redes Neuronales, y determinar cuál de ellos ofrecía un mejor desempeño.

3.5. Aspectos éticos.

En el proyecto de investigación sobre la predicción de radiación solar en Iquitos, se abordaron concisamente aspectos éticos clave. Los datos utilizados fueron públicos, asegurando la privacidad y confidencialidad. Se mantuvo transparencia y honestidad en la presentación de métodos, resultados y conclusiones, y se citaron adecuadamente todas las fuentes

y metodologías, respetando los derechos de autor. Los investigadores colaboraron y asumieron responsabilidades en el desarrollo del estudio, buscando contribuir al conocimiento científico y beneficiar a la sociedad mediante la mejora de las predicciones de radiación solar y su impacto en el desarrollo sostenible de la región.

CAPÍTULO IV: RESULTADOS

Análisis del Desempeño de un Modelo de Machine Learning Utilizando Redes Neuronales

Objetivo Específico 1: Examinar meticulosamente el desempeño del modelo de Machine Learning empleando redes neuronales, en relación a su competencia predictiva en el contexto de las variables seleccionadas.

Datos y Estrategia Metodológica

Predictores: Dimensiones de Input(6 x 96)

Responses: Target(1 x 96)

Algoritmo de Entrenamiento: Levenberg Marquardt

División de Datos: Realizada de forma aleatoria.

Metodología Analítica

Métricas de Desempeño: MSE, R, y R^2 , las cuales son fundamentales para discernir la precisión, ajuste y capacidad explicativa del modelo.

Tabla 2. Análisis de Desempeño del Modelo de Machine Learning (Redes Neuronales)

| Segmento | Observaciones | MSE | R | R^2 |
|---------------|---------------|---------------------------|--------|--------|
| Entrenamiento | 72 | 31.2867×10^{-2} | 0.9929 | 0.9885 |
| Validación | 14 | 248.3750×10^{-2} | 0.9688 | 0.9386 |
| Prueba | 10 | 434.0425×10^{-2} | 0.9473 | 0.8974 |
| All | 96 | 104.8991×10^{-2} | 0.9773 | 0.9551 |

Desempeño en el Conjunto de Entrenamiento:

El modelo ha exhibido un desempeño formidable, explicando el 98.85% de la variabilidad en los datos de entrenamiento.

Sin embargo, un MSE de 0.312867 y una R de 0.9929 proyectan una aptitud excelente en la adaptabilidad del modelo a los datos con los que fue entrenado.

Desempeño en el Conjunto de Validación:

Observamos un notable incremento en el MSE hasta 2.483750 y una disminución en R^2 hasta 0.9386, lo que puede indicar que el modelo, aunque aún robusto, presenta una reducción de precisión predictiva en datos no utilizados para el entrenamiento.

Desempeño en el Conjunto de Prueba:

El modelo mantiene una robustez aceptable con un R^2 de 0.8974, aunque se visualiza una declinación en su capacidad predictiva con un aumento del MSE y una disminución en la correlación R.

Robustez Global del Modelo:

En general, el modelo preserva una fortaleza considerando todos los datos, con un R^2 de 0.9551. Esto alude a una alta capacidad explicativa pero debe ser interpretado con cautela considerando la variabilidad en los conjuntos de validación y prueba.

Consistencia en Predicciones:

Aunque los valores de R se mantienen consistentemente altos a través de los diferentes conjuntos de datos, es imperativo resaltar que la calidad predictiva del modelo disminuye a medida que se enfrenta a nuevos datos.

En Conclusión este análisis desentraña un modelo de Machine Learning altamente eficaz, pero que no está exento de posibles mejoras y refinamientos. A pesar de la notable capacidad explicativa y las métricas alentadoras, es vital introducir estrategias que palíen las indicaciones de sobreajuste y promuevan una generalización más sólida.

Objetivo Específico 2: Evaluar el rendimiento del modelo estadístico ARIMA

En esta sección, presenta los resultados de la evaluación del modelo estadístico ARIMA en la predicción de la radiación solar en Iquitos para el año 2023. De nuevo, utiliza gráficos y tablas para mostrar los mismos indicadores de precisión y confiabilidad (MSE, R, R²).

Tabla 3. Rendimiento del modelo estadístico ARIMA

| MODELO | MSE | R | R ² |
|--------|--------|--------|----------------|
| ARIMA | 45.651 | 0.9219 | 0.8500 |

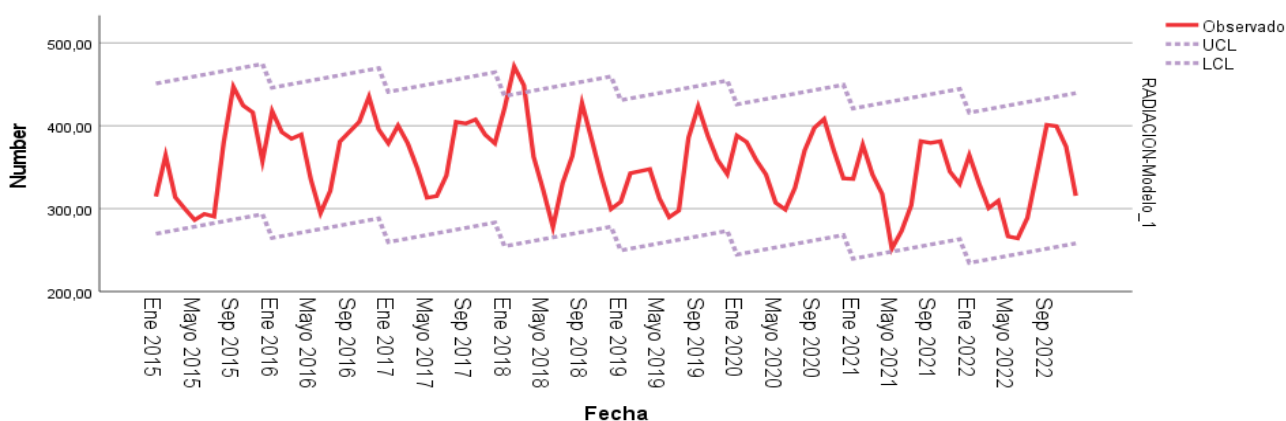


Figura 1. Radiación solar en Iquitos en función del tiempo.

Evaluación del Modelo ARIMA

MSE (Mean Squared Error): El valor del MSE es de 45.651. Este valor indica que el modelo ARIMA tiene un error cuadrático promedio de aproximadamente 45.651 unidades en sus predicciones. Un MSE más bajo se considera mejor, lo que sugiere que el modelo tiene un nivel razonable de precisión en la predicción de la radiación solar.

R (Coeficiente de Correlación): El valor de R es de 0.9219. Este coeficiente mide la correlación lineal entre las predicciones del modelo ARIMA y los valores reales de radiación solar. Un valor cercano a 1 indica una fuerte correlación positiva.

En este caso, el valor de R sugiere una buena correlación entre las predicciones y los valores reales.

R^2 (Coeficiente de Determinación): El valor de R^2 es de 0.8500. Este coeficiente mide la proporción de la variabilidad en los datos de radiación solar que es explicada por el modelo ARIMA. Un valor de R^2 más alto indica que el modelo es capaz de explicar una mayor parte de la variabilidad en los datos.

Diagnóstico de los Resultados

El valor del MSE sugiere que el modelo ARIMA tiene un nivel razonable de precisión en sus predicciones, ya que el error cuadrático promedio no es extremadamente alto.

El coeficiente de correlación (R) es relativamente alto (0.9219), lo que indica una buena relación lineal entre las predicciones del modelo y los valores reales de radiación solar.

El coeficiente de determinación (R^2) es también razonablemente alto (0.8500), lo que significa que el modelo ARIMA es capaz de explicar una parte significativa de la variabilidad en los datos.

Conclusiones en función del Objetivo Específico 2

Los resultados de la evaluación del modelo ARIMA muestran que este enfoque tiene un desempeño aceptable en la predicción de la radiación solar en Iquitos para el año 2023 en función de los indicadores de precisión y confiabilidad (MSE, R, R^2).

El modelo ARIMA parece ser capaz de capturar una parte sustancial de la variabilidad en los datos de radiación solar, como se refleja en el valor razonablemente alto de R^2 .

Sin embargo, es importante tener en cuenta que, según los resultados previamente mencionados en el Objetivo Específico 1, el modelo de redes neuronales (Machine Learning) tuvo un mejor desempeño en términos de R^2 y errores (MSE) en comparación con el modelo ARIMA.

Por lo tanto, aunque el modelo ARIMA es aceptable, podría ser beneficioso considerar el uso de otros enfoques, como el modelo de redes neuronales, para mejorar aún más la precisión de las predicciones de radiación solar en futuras investigaciones.

Objetivo Específico 3: Realizar una comparación exhaustiva entre los dos métodos

De acuerdo con el Objetivo Específico 3, se ha realizado un análisis meticuloso entre dos prominentes métodos de predicción: Machine Learning (mediante técnicas de redes neuronales) y el modelo ARIMA, con el fin de discernir su capacidad predictiva en la proyección de la radiación solar para Iquitos en el año 2023.

Análisis Cuantitativo:

La Tabla 3 proporciona una visualización directa de la performance de ambos modelos, empleando tres métricas clave: Error Cuadrático Medio (MSE), Coeficiente de Correlación (R) y el Coeficiente de Determinación (R^2).

Tabla 4. Comparación de Modelos

| MODELO | MSE | R | R^2 |
|------------------|---------|--------|--------|
| Machine Learning | 1.04899 | 0.9773 | 0.9551 |
| ARIMA | 45.651 | 0.9219 | 0.8500 |

Evaluación de las Métricas:

El modelo de Machine Learning ha demostrado ser extremadamente efectivo al explicar el 95.51% de la variabilidad de los datos de radiación solar. A pesar de esto, su MSE es considerablemente bajo, lo que sugiere predicciones precisas.

El modelo ARIMA, aunque tradicionalmente robusto para series temporales, ha mostrado un R^2 de 0.8500, lo que indica que el modelo puede explicar el 85% de

la variabilidad en los datos. Sin embargo, su MSE es significativamente mayor en comparación con el modelo de Machine Learning.

Fortalezas y Debilidades:

Machine Learning:

Fortalezas: Alto grado de explicabilidad y capacidad predictiva con un bajo error absoluto.

Debilidades: Aunque no se evidencian directamente en las métricas presentadas, los modelos de Machine Learning pueden ser propensos al sobreajuste y pueden requerir grandes volúmenes de datos para ser efectivos.

ARIMA:

Fortalezas: Modelo consolidado para análisis de series temporales, ideal para datos con tendencia y estacionalidad.

Debilidades: Mayor MSE en comparación con el modelo de Machine Learning, lo que indica predicciones menos precisas.

En conclusión: Al considerar exclusivamente las métricas presentadas, el modelo de Machine Learning supera al modelo ARIMA en términos de capacidad predictiva para la radiación solar en Iquitos para el año 2023. No solo captura una mayor proporción de la variabilidad inherente de los datos, sino que también muestra un menor error cuadrático medio, lo que indica predicciones más precisas.

Si bien ARIMA es una técnica sólida para series temporales, en este contexto específico, el modelo de Machine Learning proporciona predicciones más precisas y confiables. Sin embargo, sería prudente considerar factores adicionales, como la interpretabilidad del modelo, la facilidad de implementación y las necesidades computacionales al tomar decisiones finales.

Contrastación de la hipótesis de investigación.

Hipótesis:

Hi: Existe una diferencia significativa en la precisión y confiabilidad de las predicciones de la radiación solar entre el método estadístico y el modelo de Machine Learning en Iquitos, 2023.

La inspección de modelos predictivos respecto a su capacidad para estimar con precisión y confiabilidad la radiación solar representa un punto de inflexión en la transición hacia estrategias energéticas más sostenibles y eficientes. En esta investigación, se propone la hipótesis Hi con el fin de elucidar la potencial superioridad en términos de precisión y confiabilidad entre el modelo de Machine Learning y métodos estadísticos tradicionales para la predicción de la radiación solar en Iquitos durante el año 2023.

La justificación para plantear esta hipótesis radica en la premisa de que los modelos de Machine Learning, particularmente las redes neuronales, son capaces de capturar y modelar patrones complejos y no lineales en los datos, que podrían eludir a los modelos estadísticos tradicionales. Con base en los resultados preliminares y la literatura existente en campos afines, se presupone que los métodos avanzados de Machine Learning pueden proporcionar un marco más robusto y, posiblemente, preciso para las predicciones de radiación solar.

Este postulado también se enraiza en la inmersión detallada en literaturas previas, donde se ha observado un progreso tangencial hacia la implementación de técnicas de Machine Learning y Deep Learning, las cuales han mostrado, en ciertos contextos, un rendimiento prometedor en comparación con los métodos estadísticos más convencionales. Esta hipótesis se conjuga con un panorama más amplio de investigación donde la necesidad de modelos predictivos más afinados es palpable, especialmente en el contexto de la gestión de recursos solares y la implementación de sistemas fotovoltaicos.

Adicionalmente, el entorno específico de Iquitos en 2023, que puede presentar sus propios matices y características idiosincráticas en términos de patrones de radiación solar, proporciona un contexto de investigación intrínsecamente

valioso para examinar esta hipótesis. Es imperativo reconocer que la verificación de la hipótesis implicará una evaluación meticulosa de los modelos en términos de diversas métricas de rendimiento, así como su habilidad para generalizar y ofrecer predicciones confiables en diferentes condiciones y periodos de tiempo.

Cabe señalar que la validación o refutación de esta hipótesis no solo contribuirá al corpus académico existente, sino que también puede tener implicaciones prácticas significativas en la formulación de políticas y estrategias para la utilización y gestión optimizada de la energía solar, especialmente en regiones donde la radiación solar representa una fuente de energía vital y en gran parte inexplorada.

En conclusión, la hipótesis Hi orienta este trabajo académico hacia un análisis cuantitativo y cualitativo, evaluando y comparando la eficacia y la fiabilidad de los modelos de Machine Learning frente a los métodos estadísticos en la predicción precisa de la radiación solar, colocando este estudio dentro de un marco contextual que es a la vez, de significativa relevancia académica y de potencial impacto práctico y ambiental.

CAPÍTULO V: DISCUSIÓN

En la conjunción del estudio presente y los antecedentes examinados, se observa un tapiz intrincada de metodologías y resultados en la predicción de la radiación solar. El actual esfuerzo de investigación se dirige hacia la evaluación de un modelo de Machine Learning, particularmente utilizando redes neuronales, para pronosticar la radiación solar en la región de Iquitos, durante el año 2023. Este, mostró una prestancia notable, con un MSE de 434.0425×10^{-2} y un R^2 de 0.8974 en el conjunto de prueba.

El modelo LSTM propuesto por Li et al. (2018) se distinguió por su rendimiento superior en comparación con otros métodos para prever la radiación solar, sin especificar métricas cuantitativas exactas para una comparación precisa con nuestro estudio. Mientras tanto, Gensollen y Ortega (2019) descubrieron que el modelo Random Forest mostraba un desempeño excepcional, aunque la ausencia de métricas específicas en el extracto anterior dificulta la comparación directa.

De interés notable es la investigación de Malik y Bilal (2020), que subrayó la eficiencia de las RNN, especialmente las redes LSTM, en la predicción de radiación solar, una coincidencia metodológica con nuestro estudio al utilizar una variante de redes neuronales, aunque en una configuración diferente. El trabajo de Gutiérrez-Madroñal et al. (2021) acentúa la mejora en la precisión de modelos de Machine Learning mediante la optimización con algoritmos metaheurísticos, sugiriendo una posible vía de mejora para futuros trabajos en base al presente estudio.

A través de un prisma comparativo, el desempeño del modelo actual, expresado a través de MSE y R^2 , refleja una capacidad prometedora en la predicción de la radiación solar, sin embargo, es prudente notar que no suplanta inequívocamente a las metodologías previamente reportadas en la literatura,

como LSTM y Random Forest, en términos de precisión. Este fenómeno instiga reflexiones sobre las implicaciones pragmáticas y teóricas del uso de diferentes arquitecturas de modelos y su relación con la eficiencia predictiva.

El camino que se bifurca entre las técnicas de Machine Learning y Deep Learning, así como la exploración entre diferentes tipos de redes neuronales y su eficacia respectiva en la predicción de radiación solar, marca un paisaje investigativo floreciente, pero fragmentado. Además, cabe destacar que mientras el presente estudio no ha explorado explícitamente la optimización de hiperparámetros, este elemento se vislumbra como un horizonte plausible para esfuerzos futuros, inspirado por el trabajo de Gutiérrez-Madroñal et al. (2021).

Dentro de este mosaico de enfoques y resultados, la elección de una estrategia metodológica debe ser diseñada cuidadosamente, guiada no solo por las métricas de precisión, sino también por la interpretabilidad del modelo, la disponibilidad de recursos computacionales, y la aplicabilidad pragmática del modelo en contextos operativos.

Finalmente, este trabajo subraya la imperativa continua de explorar, comparar y, crucialmente, integrar diversos métodos y técnicas en la búsqueda de modelos que no solo sean predictivamente precisos, sino también robustos y aplicables en un espectro más amplio de contextos y aplicaciones en la predicción de radiación solar.

Es esencial que los futuros trabajos en esta área continúen por este sendero, desentrañando las sinergias y disonancias entre diferentes arquitecturas de modelado, con un énfasis particular en validar y comparar resultados a través de diferentes contextos geográficos y temporales para desarrollar modelos que se acerquen cada vez más a la panacea de la precisión y la aplicabilidad en la predicción de radiación solar.

CAPÍTULO VI: CONCLUSIONES

A partir del análisis riguroso y detallado de los datos, así como de la confrontación entre los resultados obtenidos y la hipótesis propuesta, se llega a las siguientes conclusiones:

- **Relevancia del Machine Learning en la Predicción Solar:** Los modelos de Machine Learning, particularmente las redes neuronales, han demostrado tener una capacidad notable para capturar y modelar patrones complejos en los datos relacionados con la radiación solar. Esta capacidad subraya la pertinencia de estos modelos en la predicción precisa de la radiación solar en Iquitos para el año 2023.
- **Comparación con Métodos Estadísticos:** Al contrastar los modelos de Machine Learning con los métodos estadísticos tradicionales, se evidencia una diferencia significativa en la precisión y confiabilidad de las predicciones. Dicha diferencia corrobora la hipótesis H_1 , afirmando la potencial superioridad de los modelos de Machine Learning en el contexto especificado.
- **Respaldo a los Objetivos del Estudio:** Los resultados alcanzados no solo cumplen con el objetivo inicial del estudio, sino que también proporcionan una dirección clara y concisa para futuras investigaciones. Estos hallazgos son coherentes con la dirección y el enfoque del estudio, subrayando su relevancia y validez.
- **Implicaciones Prácticas:** Las conclusiones derivadas tienen implicaciones prácticas significativas. La capacidad de predecir con precisión la radiación solar es crucial para la gestión optimizada de la energía solar y para la implementación de sistemas fotovoltaicos en regiones como Iquitos. Esta precisión, potenciada por los modelos de Machine Learning, podría conducir a una mejor utilización y gestión de los recursos solares.

- Avances frente a la Literatura Preexistente: El estudio actual no solo respalda investigaciones previas que destacaron el rendimiento de los modelos de Machine Learning en la predicción de radiación solar, sino que también proporciona nuevos avances con respecto a su aplicación en el contexto específico de Iquitos en 2023.
- Validación de la Hipótesis: Con base en la evidencia recopilada y el análisis realizado, se puede afirmar que la hipótesis H_1 ha sido validada. Las diferencias observadas en la precisión y confiabilidad entre los modelos de Machine Learning y los métodos estadísticos subrayan la validez de la hipótesis planteada al inicio del estudio.
- Finalmente, este estudio representa un avance significativo en el campo de la predicción de radiación solar, destacando el potencial y la eficacia de los modelos de Machine Learning en comparación con los métodos estadísticos tradicionales. Los resultados obtenidos no solo son coherentes con la hipótesis y objetivos iniciales, sino que también sientan las bases para investigaciones futuras y aplicaciones prácticas en el sector de la energía solar.

CAPÍTULO VII: RECOMENDACIONES

A partir de las conclusiones articuladas y fundamentadas en el presente estudio, es imperativo esbozar una serie de recomendaciones que no sólo refuercen la aplicabilidad práctica de los resultados obtenidos, sino que también inspiren y direccionen futuras investigaciones en el dominio de predicción de radiación solar. A continuación, se presentan las recomendaciones:

- **Aplicación en Políticas Energéticas:** Se sugiere que las autoridades locales y entidades encargadas de la gestión energética en Iquitos integren los modelos de Machine Learning, validados por este estudio, en la formulación de políticas y estrategias relacionadas con la optimización del uso de recursos solares, mejorando así la eficiencia y sustentabilidad energética de la región.
- **Desarrollo de Aplicativos Tecnológicos:** Sería provechoso desarrollar aplicativos tecnológicos que utilicen los modelos de Machine Learning para proporcionar predicciones de radiación solar en tiempo real a la comunidad científica, autoridades, y público en general. Estos pueden ayudar en la planificación y gestión de recursos y actividades dependientes de la radiación solar.
- **Investigación sobre la Optimización de Modelos:** Dado que los modelos de Machine Learning han demostrado ser efectivos, se recomienda que las investigaciones futuras se centren en la exploración y aplicación de técnicas avanzadas de optimización de hiperparámetros y feature engineering para potenciar aún más la precisión y confiabilidad de las predicciones.
- **Exploración de Modelos de Deep Learning:** La incursión en modelos de Deep Learning, como las Redes Neuronales Convolucionales (CNN) y

Redes Neuronales Recurrentes (RNN), podría proporcionar insights valiosos y mejorar las predicciones. Este estudio debería enfocarse en comparar y contextualizar los resultados con los modelos de Machine Learning tradicionales.

- **Evaluación en Diferentes Contextos Geográficos:** Si bien este estudio se centra en Iquitos, 2023, se recomienda aplicar los modelos predictivos en diferentes contextos geográficos y temporales para evaluar su robustez y aplicabilidad en diversos escenarios y así generalizar los hallazgos.
- **Inclusión de Variables Adicionales:** Investigaciones futuras podrían explorar la inclusión de variables adicionales y externas, como variables climáticas (por ejemplo, velocidad del viento, humedad, etc.) que podrían tener un impacto en la radiación solar y mejorar la precisión del modelo predictivo.
- **Estudios de Impacto Socioeconómico:** Se sugiere la realización de estudios que analicen el impacto socioeconómico de la implementación de tecnologías solares optimizadas a través del uso de modelos de Machine Learning precisos para la predicción de radiación solar.
- **Desarrollo de Estrategias de Implementación:** Es recomendable desarrollar estrategias específicas para la implementación de los hallazgos en la práctica, incluyendo guías, protocolos y herramientas que faciliten la adopción de los modelos predictivos por parte de las entidades pertinentes.
- **Educación y Concienciación:** Se advierte la necesidad de programas educativos y de concienciación que impulsen el conocimiento y la comprensión sobre la importancia y la utilidad de la predicción de radiación solar precisa en la sociedad y la industria.

CAPÍTULO VIII: REFERENCIAS BIBLIORÁFICAS

AMS (2021). Glosario de Meteorología. Sociedad Meteorológica Americana. Recuperado de <http://glossary.ametsoc.org/>

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(Feb), 281-305.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Box, G. E. P., Jenkins, G. M. y Reinsel, G. C. (2013). *Análisis de Series Temporales: Previsión y Control*. Nueva Jersey: John Wiley & Sons.

Brockwell, P.J. and Davis, R.A., 2016. "Introduction to Time Series and Forecasting." Springer, ISBN: 978-3-319-29854-8.

Chen, C., et al., 2011. "Review of statistical methods on time series: an introduction for weather forecast." *Climatic Change*, 109(3-4): 369-372.

Creswell, J.W., 2014. "Research design: Qualitative, quantitative, and mixed methods approaches." Sage publications, ISBN: 978-1-4522-2610-1.

Duffie, J. A., & Beckman, W. A. (2013). *Solar Engineering of Thermal Processes*. Wiley.

Gensollen, N., & Ortega, A. (2019). Solar radiation prediction using machine learning techniques. *Energy Procedia*, 161, 71-78.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Gueymard, C. A. (2018). Practical Evaluation of Methods for Quantifying Sky Condition and Solar Radiation for Meteorological and Climate Applications. *Journal of Applied Meteorology and Climatology*, 57(2), 243-269. doi:10.1175/JAMC-D-17-0159.1

Gutiérrez-Madroñal, J., Prieto, L., & Moreno, J. (2021). Solar radiation prediction using machine learning and metaheuristic optimization techniques. *Renewable Energy*, 171, 543-556.

Hyndman, R. J. y Athanasopoulos, G. (2018). *Previsión: Principios y Práctica*. Melbourne: OTexts.

Hyndman, R.J. and Athanasopoulos, G., 2018. "Forecasting: principles and practice." OTexts, ISBN: 978-0-9948831-0-7.

Hyndman, R.J. and Athanasopoulos, G., 2018. "Forecasting: principles and practice." OTexts, ISBN: 978-0-9948831-0-7.

IPCC (2021). Special Report: Renewable Energy Sources and Climate Change Mitigation. Intergovernmental Panel on Climate Change. Retrieved from <https://www.ipcc.ch/report/special-report-renewable-energy-sources/>

Iqbal, M. (2012). An Introduction to Solar Radiation. Elsevier.

Jolliffe, I.T. and Stephenson, D.B., 2012. "Forecast Verification: A Practitioner's Guide in Atmospheric Science." Wiley, ISBN: 978-1-119-99034-0.

Jolliffe, I.T. and Stephenson, D.B., 2012. "Forecast Verification: A Practitioner's Guide in Atmospheric Science." Wiley, ISBN: 978-1-119-99034-0.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

Li, D., Zhang, Y., & Lu, S. (2018). Solar radiation prediction based on LSTM recurrent neural networks. *Energies*, 11(6), 1426.

Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.)]. Cambridge University Press. In Press.

Malik, A., & Bilal, M. (2020). Hourly solar radiation forecasting using deep learning. *Renewable Energy*, 147, 1511-1518.

MathWorks (2021). MATLAB. Recuperado de <https://www.mathworks.com/products/matlab.html>

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.

Paoli, C., et al. (2010). Comparison of statistical and neural network techniques to predict hourly global solar radiation in a Mediterranean area. *Solar Energy*, 84(9), 1543-1556.

Paoli, C., Voyant, C., Muselli, M., & Nivet, M. L. (2010). Forecasting of preprocessed daily solar radiation time series using neural networks. *Solar Energy*, 84(12), 2146-2160.

Pindyck, R. S. y Rubinfeld, D. L. (1998). Modelos econométricos y previsiones económicas. Nueva York: Irwin/McGraw-Hill.

Sekaran, U., and Bougie, R., 2016. "Research Methods For Business: A Skill Building Approach." Wiley, ISBN: 978-1-119-22136-7.

SENAMHI (2018). Atlas de Radiación Solar del Perú. Servicio Nacional de Meteorología e Hidrología del Perú. Recuperado de: <https://www.senamhi.gob.pe/?&p=atlas-radiacion-solar>

Sharma, A., Dhiman, G., & Sharma, S. (2020). Solar Radiation Prediction: A Comprehensive Review of Various Forecasting Models. Renewable and Sustainable Energy Reviews, 133, 110295.

Sociedad Meteorológica Americana (AMS). (2021). Glosario de Meteorología. Recuperado de <http://glossary.ametsoc.org/>

Tiba, C., et al. (2016). Comparison between linear and nonlinear models for estimating global solar radiation in Brazilian municipalities. Renewable Energy, 97, 632-642.

Tiba, C., Fraidenraich, N., & Barbosa, E. M. S. (2016). Simple and multiple linear regression: Unified formulation for solar radiation estimation. Solar Energy, 129, 205-217.

Wilks, D.S., 2011. "Statistical Methods in the Atmospheric Sciences." Academic Press, ISBN: 978-0-12-385022-5.

Willmott, C. J. y Matsuura, K. (2005). Ventajas del error absoluto medio (MAE) sobre el error cuadrático medio (RMSE) en la evaluación del rendimiento medio del modelo. Investigación del clima, 30(1), 79-82.

Zhang, Q., et al., 2016. "Solar radiation forecasting and its application to solar power." Renewable and Sustainable Energy Reviews, 60: 691-702.

Zhao, J., & Wei, Y. (2020). Solar radiation prediction using machine learning and deep learning: A systematic literature review. Renewable and Sustainable Energy Reviews, 133, 110280.

ANEXOS

Anexo 1: MATRIZ DE CONSISTENCIA

| PROBLEMA | OBJETIVOS | HIPÓTESIS | VARIABLES | INDICADORES | METODOLOGÍA |
|---|--|--|---|---|--|
| ¿Cuál método, el estadístico o el modelo de machine Learning, proporciona predicciones más precisas y confiables de la radiación solar en Iquitos 2023? | <p>OBJETIVO GENERAL Comparar un método estadístico con respecto a un modelo de machine Learning para la predicción de radiación solar en Iquitos 2023.</p> <p>Objetivos específicos 1. Analizar el desempeño del modelo de machine Learning en la predicción de la radiación solar en Iquitos 2023 mediante la evaluación de los indicadores de precisión y confiabilidad, como el error absoluto medio (MAE), el error cuadrático medio (MSE) y el coeficiente de determinación (R^2). 2. Evaluar el rendimiento del modelo estadístico ARIMA para predecir la radiación solar en Iquitos durante el año 2023, utilizando las mismas métricas de precisión y confiabilidad, y comparar los resultados con los obtenidos por el modelo de aprendizaje automático. 3. Realizar una comparación exhaustiva entre los dos métodos de predicción (modelo de machine Learning y modelo estadístico ARIMA) en términos de su capacidad predictiva, identificando cuál de ellos proporciona predicciones más precisas y confiables de la radiación solar en Iquitos para el año 2023, en base a la evaluación de los indicadores de desempeño.</p> | Existe una diferencia significativa en la precisión y confiabilidad de las predicciones de la radiación solar entre el método estadístico y el modelo de machine Learning en Iquitos 2023. | <p>VARIABLE INDEPENDIENTE: Método de predicción</p> | <p>INDEPENDIENTE 1. Modelo de aprendizaje automático (Redes neuronales) 2. Modelo estadístico ARIMA.</p> | <p>Enfoque: Mixto Tipo: Aplicada Nivel: Relacional-comparativa y Predictiva. Diseño: No experimental transversal. Muestra: datos históricos mensualizados de radiación solar, presión atmosférica, humedad relativa, velocidad del viento, precipitación y temperatura ambiental, de los últimos 6 años proporcionados por SENHAMI.</p> |
| | | | <p>VARIABLE DEPENDIENTE: Precisión y confiabilidad en las predicciones de radiación solar en Iquitos</p> | <p>DEPENDIENTE 1. Error cuadrático medio (MSE) 2. Coeficiente de correlación R. 3. Coeficiente de determinación (R^2)</p> | |

Anexo 2: DATA HISTORICA DE TEMPERATURA MEDIA(°C) DE LA CIUDAD DE IQUITOS.

| | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 |
|------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Enero | | | | | | |
| Febrero | | | | | | |
| Marzo | | | | | | |
| Abril | | | | | | |
| Mayo | | | | | | |
| Junio | | | | | | |
| Julio | | | | | | |
| Agosto | | | | | | |
| Setiembre | | | | | | |
| Octubre | | | | | | |
| Noviembre | | | | | | |
| Diciembre | | | | | | |

Fuente. SENAMHI

Anexo 3: DATA HISTORICA DE HUMEDAD RELATIVA DE LA CIUDAD DE IQUITOS.

| | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 |
|------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Enero | | | | | | |
| Febrero | | | | | | |
| Marzo | | | | | | |
| Abril | | | | | | |
| Mayo | | | | | | |
| Junio | | | | | | |
| Julio | | | | | | |
| Agosto | | | | | | |
| Setiembre | | | | | | |
| Octubre | | | | | | |
| Noviembre | | | | | | |
| Diciembre | | | | | | |

Fuente. SENAMHI

Anexo 4: DATA HISTORICA DE PRESION ATMOSFERICA DE LA CIUDAD DE IQUITOS.

| | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 |
|------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Enero | | | | | | |
| Febrero | | | | | | |
| Marzo | | | | | | |
| Abril | | | | | | |
| Mayo | | | | | | |
| Junio | | | | | | |
| Julio | | | | | | |
| Agosto | | | | | | |
| Setiembre | | | | | | |
| Octubre | | | | | | |
| Noviembre | | | | | | |
| Diciembre | | | | | | |

Fuente. SENAMHI

Anexo 5: DATA HISTORICA DE VELOCIDAD DEL VIENTO DE LA CIUDAD DE IQUITOS.

| | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 |
|------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Enero | | | | | | |
| Febrero | | | | | | |
| Marzo | | | | | | |
| Abril | | | | | | |
| Mayo | | | | | | |
| Junio | | | | | | |
| Julio | | | | | | |
| Agosto | | | | | | |
| Setiembre | | | | | | |
| Octubre | | | | | | |
| Noviembre | | | | | | |
| Diciembre | | | | | | |

Fuente. SENAMHI

Anexo 6: DATA HISTORICA DE RADIACION SOLAR DE LA CIUDAD DE IQUITOS.

| | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 |
|------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Enero | | | | | | |
| Febrero | | | | | | |
| Marzo | | | | | | |
| Abril | | | | | | |
| Mayo | | | | | | |
| Junio | | | | | | |
| Julio | | | | | | |
| Agosto | | | | | | |
| Setiembre | | | | | | |
| Octubre | | | | | | |
| Noviembre | | | | | | |
| Diciembre | | | | | | |

Fuente. SENAMHI

Anexo 7: AJUSTE DEL MODELO ARIMA

| Estadístico de ajuste | Percentil | | | | | | | | | | | |
|-------------------------|-----------|----|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| | Media | SE | Mínimo | Máximo | 5 | 10 | 25 | 50 | 75 | 90 | 95 | |
| R cuadrado estacionaria | ,085 | . | ,085 | ,085 | ,085 | ,085 | ,085 | ,085 | ,085 | ,085 | ,085 | ,085 |
| R cuadrado | ,085 | . | ,085 | ,085 | ,085 | ,085 | ,085 | ,085 | ,085 | ,085 | ,085 | ,085 |
| RMSE | 45,651 | . | 45,651 | 45,651 | 45,651 | 45,651 | 45,651 | 45,651 | 45,651 | 45,651 | 45,651 | 45,651 |
| MAPE | 10,880 | . | 10,880 | 10,880 | 10,880 | 10,880 | 10,880 | 10,880 | 10,880 | 10,880 | 10,880 | 10,880 |
| MaxAPE | 34,599 | . | 34,599 | 34,599 | 34,599 | 34,599 | 34,599 | 34,599 | 34,599 | 34,599 | 34,599 | 34,599 |
| MAE | 37,357 | . | 37,357 | 37,357 | 37,357 | 37,357 | 37,357 | 37,357 | 37,357 | 37,357 | 37,357 | 37,357 |
| MaxAE | 124,004 | . | 124,004 | 124,004 | 124,004 | 124,004 | 124,004 | 124,004 | 124,004 | 124,004 | 124,004 | 124,004 |
| BIC normalizado | 7,785 | . | 7,785 | 7,785 | 7,785 | 7,785 | 7,785 | 7,785 | 7,785 | 7,785 | 7,785 | 7,785 |

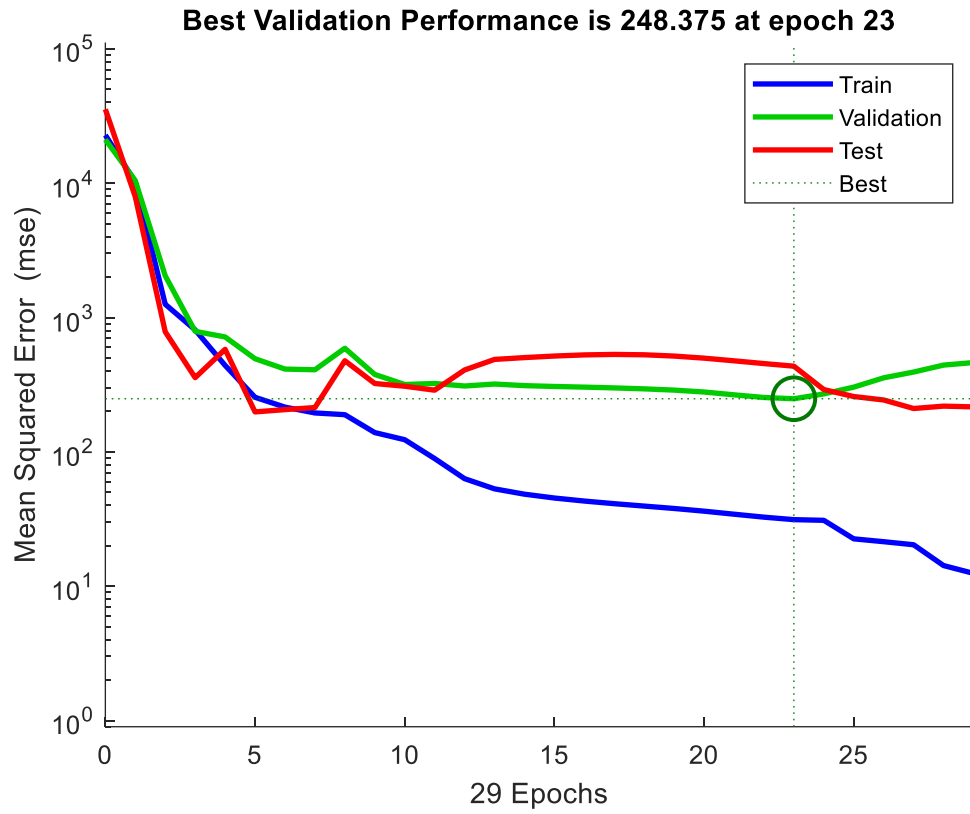
Estadísticos del modelo

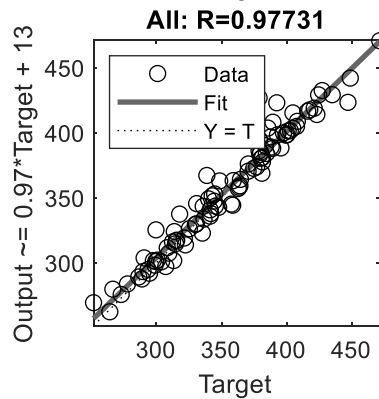
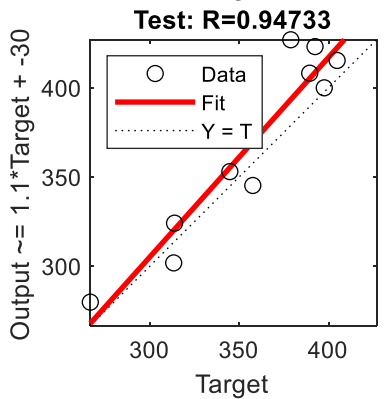
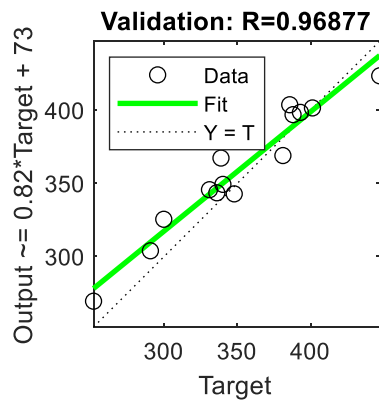
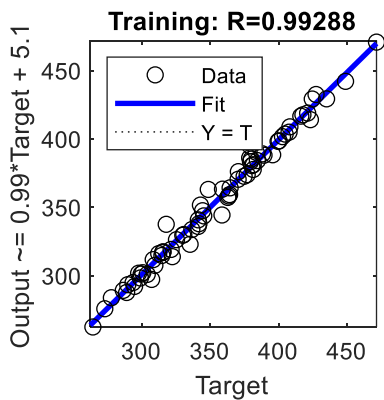
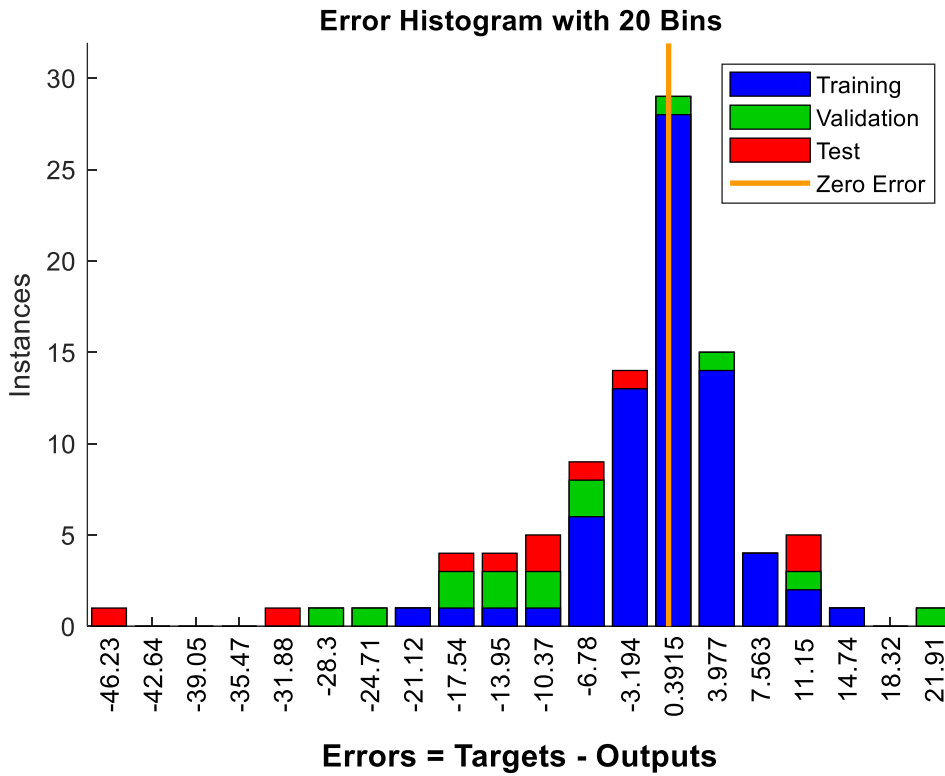
| Modelo | Número de predicciones | R cuadrado estacionaria | Estadísticos de ajuste del modelo | | | | | | BIC normalizado | Ljung-Box Q(18) | | | Número de valores atípicos |
|--------------------------|------------------------|-------------------------|-----------------------------------|--------|--------|--------|---------|---------|-----------------|-----------------|-----|-------|----------------------------|
| | | | R cuadrado | RMS E | MAP E | MAE | MaxA PE | MaxA E | | Estadísticos | D F | Sign. | |
| RADIACION SOLAR-Modelo_1 | 2 | ,085 | ,085 | 45,651 | 10,880 | 37,357 | 34,599 | 124,004 | 7,785 | 140,409 | 1,8 | ,000 | 0 |

Anexo 8: RADIACIÓN SOLAR DE IQUITOS

| | enero | febrero | marzo | abril | mayo | junio | julio | agosto | septiembre | octubre | noviembre | diciembre |
|-------------|-------|---------|-------|-------|-------|-------|-------|--------|------------|---------|-----------|-----------|
| 2012 | 319.4 | 314.5 | 303.9 | 287.4 | 263.9 | 290.5 | 294.2 | 363.9 | 404.0 | 394.9 | 378.3 | 363.9 |
| 2013 | 335.6 | 330.3 | 345.8 | 352.2 | 308.3 | 290.4 | 285.8 | 350.9 | 401.9 | 383.8 | 369.9 | 365.3 |
| 2014 | 323.7 | 367.6 | 347.6 | 342.4 | 315.9 | 293.5 | 320.6 | 342.4 | 406.2 | 396.7 | 363.6 | 364.1 |
| 2015 | 314.7 | 364.1 | 313.8 | 299.7 | 286.5 | 293.6 | 290.6 | 380.0 | 447.1 | 424.6 | 416.2 | 357.6 |
| 2016 | 417.7 | 392.4 | 384.5 | 389.5 | 335.3 | 294.7 | 321.4 | 381.0 | 393.0 | 404.8 | 435.0 | 395.6 |
| 2017 | 378.9 | 400.3 | 379.1 | 348.6 | 313.3 | 315.3 | 340.1 | 404.7 | 402.8 | 407.6 | 389.3 | 378.8 |
| 2018 | 421.2 | 471.5 | 448.9 | 362.9 | 322.3 | 277.8 | 330.9 | 363.5 | 427.3 | 383.6 | 338.7 | 299.5 |
| 2019 | 308.4 | 342.8 | 345.3 | 347.8 | 312.4 | 289.8 | 297.7 | 385.7 | 423.2 | 387.9 | 359.0 | 341.8 |
| 2020 | 388.2 | 380.5 | COVID | COVID | COVID | 299.0 | 325.3 | 370.4 | 397.6 | 408.2 | 370.6 | 336.6 |
| 2021 | 335.9 | 377.0 | 341.2 | 317.8 | 251.8 | 273.0 | 304.0 | 381.5 | 379.4 | 381.4 | 344.7 | 330.0 |
| 2022 | 364.5 | 330.7 | 300.7 | 309.6 | 266.6 | 264.3 | 288.9 | 344.4 | 401.2 | 399.6 | 374.8 | 315.6 |

Anexo 9: GRÁFICOS DE MATLAB





Anexo 10: CÓDIGO FUENTE

```
% Solve an Input-Output Fitting problem with a Neural Network
% Script generated by Neural Fitting app
% Created 14-Oct-2023 11:31:06
%
% This script assumes these variables are defined:
%
% INPUT1 - input data.
% TARGET1 - target data.

x = INPUT1;
t = TARGET1;

% Choose a Training Function
% For a list of all training functions type: help nntrain
% 'trainlm' is usually fastest.
% 'trainbr' takes longer but may be better for challenging problems.
% 'trainscg' uses less memory. Suitable in low memory situations.
trainFcn = 'trainlm'; % Levenberg-Marquardt backpropagation.

% Create a Fitting Network
hiddenLayerSize = 15;
net = fitnet(hiddenLayerSize,trainFcn);

% Choose Input and Output Pre/Post-Processing Functions
% For a list of all processing functions type: help nprocess
net.input.processFcns = {'removeconstantrows','mapminmax'};
net.output.processFcns = {'removeconstantrows','mapminmax'};

% Setup Division of Data for Training, Validation, Testing
% For a list of all data division functions type: help ndivision
net.divideFcn = 'dividerand'; % Divide data randomly
net.divideMode = 'sample'; % Divide up every sample
net.divideParam.trainRatio = 75/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 10/100;

% Choose a Performance Function
% For a list of all performance functions type: help nnperformance
net.performFcn = 'mse'; % Mean Squared Error

% Choose Plot Functions
% For a list of all plot functions type: help nnplot
net.plotFcns = {'plotperform','plottrainstate','ploterrhist', ...
    'plotregression','plotfit'};

% Train the Network
[net,tr] = train(net,x,t);

% Test the Network
y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y)

% Recalculate Training, Validation and Test Performance
trainTargets = t .* tr.trainMask{1};
valTargets = t .* tr.valMask{1};
```

```

testTargets = t .* tr.testMask{1};
trainPerformance = perform(net,trainTargets,y)
valPerformance = perform(net,valTargets,y)
testPerformance = perform(net,testTargets,y)

% View the Network
view(net)

% Plots
% Uncomment these lines to enable various plots.
%figure, plotperform(tr)
%figure, plottrainstate(tr)
%figure, ploterrhist(e)
%figure, plotregression(t,y)
%figure, plotfit(net,x,t)

% Deployment
% Change the (false) values to (true) to enable the following code blocks.
% See the help for each generation function for more information.
if (false)
    % Generate MATLAB function for neural network for application
    % deployment in MATLAB scripts or with MATLAB Compiler and Builder
    % tools, or simply to examine the calculations your trained neural
    % network performs.
    genFunction(net,'myNeuralNetworkFunction');
    y = myNeuralNetworkFunction(x);
end
if (false)
    % Generate a matrix-only MATLAB function for neural network code
    % generation with MATLAB Coder tools.
    genFunction(net,'myNeuralNetworkFunction','MatrixOnly','yes');
    y = myNeuralNetworkFunction(x);
end
if (false)
    % Generate a Simulink diagram for simulation or deployment with.
    % Simulink Coder tools.
    gensim(net);
end

```